(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 438 174 A2**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: **91100591.6**

(51) Int. Cl.5: **G10L 3/00**

(22) Date of filing: **18.01.91**

(30) Priority: **18.01.90 JP 8593/90**
18.01.90 JP 8594/90
13.02.90 JP 33209/90
13.02.90 JP 33212/90

(43) Date of publication of application:
24.07.91 Bulletin 91/30

(84) Designated Contracting States:
CH DE FR GB LI NL SE

(71) Applicant: **MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.**
1006, Oaza Kadoma
Kadoma-shi, Osaka-fu, 571(JP)

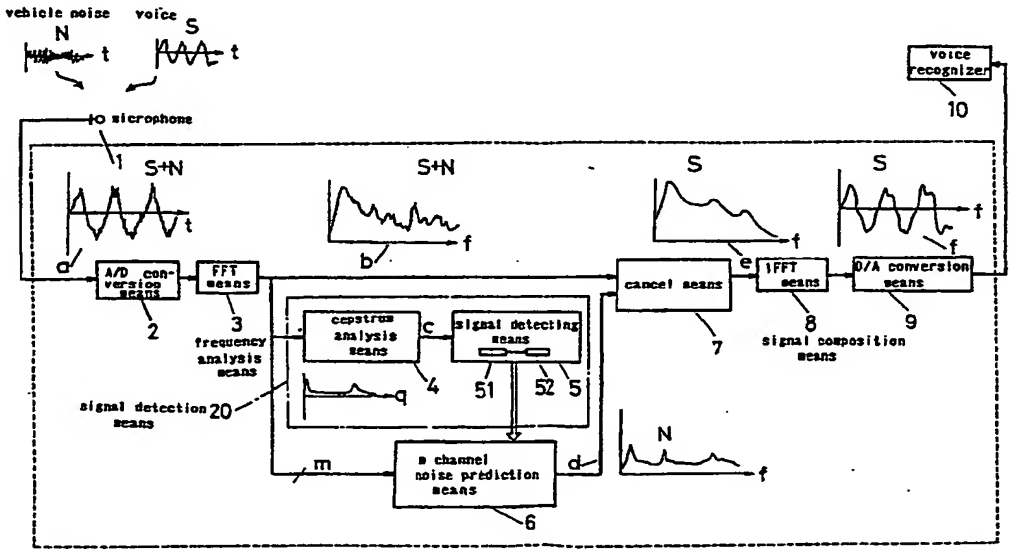
(72) Inventor: **Kane, Joji**
4-34, Jingu 1-chome
Nara-shi Nara 631(JP)
Inventor: **Nohara, Akira**
10-17, Matsuzonocho
Nishinomiya-shi Hyogo 662(JP)

(74) Representative: **Patentanwälte Beetz sen. - Beetz jun. Timpe - Siegfried - Schmitt-Fumian- Mayr**
Steinsdorfstrasse 10
W-8000 München 22(DE)

(54) Signal processing device.

(57) Frequency analysis means (3) inputs therein a signal containing a noise to perform frequency analysis; signal detection means (20) detect a signal portion from the frequency-analyzed signal ; noise prediction means (6) input therein the frequency-analyzed signal with a mingling of noise to predict a noise of the signal portion on the basis of the past information; cancel means (7) subtract the predicted noise from said frequency-analyzed signal with a mingling of noise; and band composition means (8) perform band composition on the canceled signal.

FIG. 2



EP 0 438 174 A2

BEST AVAILABLE COPY

BNSDOCID: <EP_____0438174A2_I_>

SIGNAL PROCESSING DEVICE

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a signal processing device for eliminating effectively a noise from a signal containing a noise such as a signal with a mingling of noise.

2. Description of the Related Art

Fig. 1 is a diagram showing the outline of a prior art noise suppression system (Japanese Patent Application Publication No. 63-500543).

In Fig. 1, a voice-plus-noise signal at an input is divided by a channel divider 19 into many selected channels. Then, the gain of these individual pre-processed voice channels is adjusted by a channel gain modifier 21 in response to a modified signal described later so that the gain of the channels exhibiting a low voice-to-noise ratio is reduced. Then, the individual channels comprising the post-processed voice are recombined in a channel combiner 26 to form a noise-suppressed voice signal available at an output.

Also, the individual channels comprising the pre-processed voice are applied to a channel energy estimator 20 which serves to generate energy envelope values for each channel. The post-processed voice is inputted into a channel energy estimator 22. The post-processed estimated channel energy is utilized by a background noise estimator 23 to determine voice/noise.

A channel SNR estimator 24 compares the background noise estimate of the estimator 23 to the channel energy estimate of the estimator 20 to form an SNR estimate. The SNR estimate is utilized to select a specified gain value from a channel gain table comprising experimentally beforehand determined gains. A channel gain controller 25 generates the individual channel gain values of the modified signal in response to the SNR estimate.

However, such prior art noise suppression system adjusts the gain to suppress noise without eliminating noise, so that a sufficiently noise-eliminated signal cannot be obtained.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a signal processing device for generating a securely noise-eliminated signal by solving the problem of such prior art noise suppression system.

A signal processing device of claim 1 comprises:
frequency analysis means for inputting therein a signal containing a noise to perform frequency analysis;
signal detection means for detecting a signal portion from the frequency-analyzed signal ;
noise prediction means for inputting therein the frequency-analyzed signal to predict the noise of the signal portion on the basis of a past noise information;
cancel means for subtracting the predicted noise from the frequency-analyzed signal ; and
signal composition means for performing signal composition on the noise-canceled signal.

A signal processing devise of claim 6 comprises:
frequency analysis means for analyzing the frequency of an inputted signal containing a noise;
noise prediction means for inputting therein the frequency-analyzed output of the frequency analysis means to predict the noise component thereof;
cancel means for inputting therein the predicted noise of the noise prediction means, the frequency-analyzed output of the frequency analysis means and a cancel coefficient signal to cancel the noise component considering the cancel coefficient from the frequency-analyzed output; and
signal composition means for composing the canceled-output of the cancel means.

A signal processing device of claim 7 comprises:
band division means for dividing the frequency band of an inputted signal containing a noise;
noise prediction means for inputting therein the band-divided output of the band division means to predict the noise component;
cancel means for inputting therein the predicted noise output of the noise prediction means, the band-divided output of the band division means and a cancel coefficient signal to cancel the noise component considering the cancel coefficient from the band-divided output; and
band composition means for composing the canceled-output of the cancel means.

BNSDOCID: <EP 0438174A2 1 >

A signal processing device of claim 9 comprises:
a noise prediction section for outputting a predicted noise signal by a voice/noise signal input;
a cancel section for inputting therein the voice/ noise signal input and the predicted noise signal to obtain a noise-eliminated voice output;
a voice detection section for detecting the presence/absence of a voice signal from the noise-canceled voice signal to output a control signal corresponding to the presence/absence; and
the noise prediction section performing noise prediction utilizing the control signal of the voice detection section.

A signal processing device of claim 10 comprises:
a first noise prediction section for outputting a first predicted noise signal by a voice/noise signal input;
a first cancel section into which the voice/noise signal and the first predicted noise signal are inputted, to obtain a first noise-eliminated voice output;
a voice detection section for detecting the presence/absence of a voice signal from the noise-canceled voice signal to output a control signal corresponding to the presence/absence;
the first noise prediction section performing noise prediction utilizing the control signal of the voice detection section; and further
the device provided with ;
a second noise prediction section for outputting a second predicted noise signal by the first voice output and by the control signal;
a second cancel section into which the first noise-eliminated output and the second predicted noise signal are inputted to obtain a second noise-eliminated voice output.

A signal processing device of claim 11 comprises:
FFT processing means for FFT processing a signal containing a noise;
cepstrum peak detection means for performing cepstrum analysis of the FFT output of the FFT processing means to detect the peak thereof;
pitch frequency estimate means for estimating a pitch frequency in the peak-detected output of the cepstrum peak detection means;
window generation means for generating a window output on the basis of the pitch frequency predicted output in the pitch frequency estimate means;
noise prediction means for predicting a noise from the FFT output of the FFT processing means;
cancel means for canceling the predicted noise output of the noise prediction means from the FFT output of the FFT processing means;
pitch frequency emphasis means for emphasizes the canceled output of the cancel means by the window output of the window generation means; and
IFFT processing means for IFFT processing the emphasized output of the pitch frequency emphasis means

BRIEF DESCRIPTION OF THE DRAWING

Fig. 1 is a block diagram showing a prior art noise suppression system;
Fig. 2 is a block diagram showing an embodiment of a signal processing device according to the present invention;
Fig. 3 is graphs of spectrum and cepstrum in the embodiment;
Fig. 4 is a graph illustrating a noise prediction method in the embodiment;
Fig. 5 is a graph illustrating a cancellation method with the time as a basis in the embodiment;
Fig. 6 is a graph illustrating a cancellation method with the frequency as a basis in the embodiment;
Fig. 7 is a block diagram showing an embodiment of a signal processing device according to another present invention;
Fig. 8 is a block diagram showing an embodiment of a signal processing device according to another present invention;
Fig. 9 is a graph illustrating a cancel coefficient in the embodiment;
Fig. 10 is a block diagram showing an embodiment of a signal processing device according to another present invention;
Fig. 11 is a block diagram showing an embodiment of a signal processing device according to another present invention; and
Fig. 12 is a block diagram showing an embodiment of a signal processing device according to another present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

With reference to the drawings, embodiments of the present invention will be explained hereinafter.

BNSDOCID: <EP 0438174A2 I >

Fig. 2 is a block diagram showing an embodiment of a signal processing device according to the present invention.

Into a microphone 1, usually a noise such as engine sound in addition to voice S is entered. Accordingly, the microphone 1 outputs a voice signal with a mingling of noise(S + N).

A/D (Analog-to-Digital) conversion means 2 converts the voice signal with a mingling of noise being an analog signal to a digital signal.

FFT (Fast Fourier Transformation) means 3 as an example of frequency analysis means performs fast Fourier transformation on the voice signal with a mingling of noise thus converted to the digital signal.

Signal detection means 45 detects a signal portion from the signal with a mingling of noise thus Fourier-transformed. For example, the means 45 is provided with a cepstrum analysis means 4 for cepstrum analyzing the Fourier-transformed signal and signal detecting means 5 for detecting the signal portion utilizing the cepstrum thus analyzed. The term "cepstrum" which is derived from the term "spectrum" is in this application symbolized by $c(\tau)$ and obtained by inverse-Fourier-transforming the logarithm of a short-time spectrum $S(\omega)$.

$$c(\tau) = \sum_{m=0}^{M} \log |S(\omega_m)|^2 \cos(\tau \omega_m)$$

The dimension of τ is time and τ (time) is named "quefrency" which is derived from the word "frequency".

As example the waveforms as shown in Fig. 3 are given. Fig. 3(a) is a short-time spectrum, and Fig. 3-(b) is a cepstrum thereof. The signal detecting means 5 detects the signal portion from a noise portion utilizing the cepstrum. As a method of discriminating the signal portion utilizing the cepstrum, for example, a method has been known of detecting the peak of the cepstrum. That is, the method utilizes a peak detection means 51 for detecting the peak of the analyzed cepstrum and signal-noise detection means 52 for discriminating signal on the basis of the peak information thus detected. The P in Fig. 3(b) shows the peak, and the portion in which the peak exists is determined to be a voice signal portion. The peak is detected, for example, in such a manner that a specified threshold has been previously set and the peak value is compared to the threshold.

Noise prediction means 6 inputs therein the Fourier-transformed signal with a mingling of noise and predicts the noise in the signal portion on the basis of a past noise information. For example, as shown in Fig. 4, the axis X represents frequency, the axis Y does voice level and the axis Z does time. The data of p1 and p2 through pi at a frequency fl are taken to predict the preceding pj. For example, the mean value of the noise portions p1 through pi is predicted to be pj. Alternatively, when the voice signal portion continues further, the pj is further multiplied by an attenuation coefficient. In the present embodiment, the noise prediction means 6 predicts the noise in the signal portion utilizing the signal portion information detected by the signal detection means 45. For example when the signal portion is detected, the means 45 predicts the noise in the signal portion on the basis of the data of the noise portion at the nearest past when viewed from the point beginning with the signal portion. It is also preferable that the noise prediction means 6 utilizes the signal portion (noise portion) information detected by the signal detection means 45 to accumulate the past noise information.

Cancel means 7 subtracts the noise predicted by the noise prediction means 6 from the Fourier-transformed signal with a mingling of noise. For example, the cancellation with the time as a basis is performed in a manner to subtract the predicted noise waveform (b) from the noise-contained voice signal (a) as shown in Fig. 5, thereby allowing only the signal to be taken out (c). Also, as shown in Fig. 6, the cancellation with the frequency as a basis is performed in such a manner that the noise-contained voice signal (a) is Fourier transformed (b), then from the signal thus transformed ,the predicted noise spectrum (c) is subtracted (d), and the remain is inverse Fourier transformed to obtain a voice signal without noise (e). Of course, the portion without signal can be determined to be only noise, so that a signal obtained by inversing the output of the FFT means 3 is generated, and in the portion without signal, the inversed signal is added directly to the output of the FFT means 3 to eliminate completely noise.

IFFT means 8, as an example of signal composition means, inverse-Fourier transforms the noise-eliminated signal obtained by the cancel means 7.

D/A conversion means 9 converts the noise-eliminated voice signal being a digital signal obtained by the IFFT means 8 to an analog signal. The f in FIG.2 indicates the noise-eliminated signal being the analog signal.

A voice recognizer 10 recognizes what word the noise-eliminated voice signal thus obtained is.

The operation of the above-described embodiment will be explained hereinafter.

The microphone 1 inputs therein a voice with a mingling of noise and outputs the voice signal with a mingling of noise (S + N) (see Fig. 2, a). The A/D conversion means 2 converts the voice signal with a mingling of noise being an analog signal to a digital signal. The FFT means 3 performs fast Fourier transformation on the voice signal with a mingling of noise thus converted to the digital signal (see Fig.2, b).

The signal detection means 45 detects a signal portion from the signal with a mingling of noise thus Fourier transformed. For example, the cepstrum analysis means 4 performs cepstrum analysis on the Fourier-transformed signal. Further, the signal detection means 5 detects the signal portion utilizing the cepstrum thus analyzed (see Fig.2, c). For example the means 5 detects the peak of the cepstrum to detect signal .

The noise prediction means 6 inputs therein the Fourier-transformed signal with a mingling of noise, takes the data of p1 and p2 through pi at a frequency fl, and calculates the mean value of the noise portions p1 through pi to be made pj. Also, in the present embodiment, the noise prediction means 6 predicts the noise in the signal portion (see Fig. 2, d) , on the basis of the data of the noise portion at the nearest past when viewed from the point beginning with the signal portion when the signal is detected utilizing the signal portion information detected by the signal detection means 45 .

The cancel means 7 subtracts the noise predicted by the noise prediction means 6 from the Fourier-transformed signal with a mingling of noise (see Fig. 2, e).

The IFFT means 8 inverse-Fourier transforms the noise-eliminated signal obtained by the cancel means 7.

The D/A conversion means 9 converts the noise-eliminated voice signal being a digital signal obtained by the IFFT means 8 to an analog signal (see Fig.2, f).

The voice recognizer 10 recognizes what word the noise-eliminated voice signal thus obtained is. Since the signal contains no noise, the recognition rate thereof becomes high.

The noise detection means 6 of the present invention may be such means as to predict the noise component of the signal portion simply on the basis of the past noise information without utilizing the detected signal from the signal detection means 45. For example, the means 6 predicts simply that the past noise continues even in the signal portion.

The present invention also can apply to the processing of other signals with a mingling of noise not limiting to that of voice signal.

Further, the present invention, though implemented in software utilizing a computer, may also be implemented utilizing a dedicated hardware circuit.

As described above, the signal processing device according to the present invention detects a signal portion from a frequency-analyzed signal with a mingling of noise, predicts a noise of the signal portion on the basis of the past noise information, and subtracts the predicted noise from the signal with a mingling of noise, thereby allowing a completely noise-eliminated signal to be generated.

Where noise prediction means 6 uses a signal detected by signal detection means 45 as a trigger to predict a noise of the signal portion, the noise can more accurately predicted, whereby a signal from which the noise is more securely eliminated can be generated.

With reference to the drawings, embodiments of another present invention will be explained hereinafter.

Fig. 7 is a block diagram of a signal processing device in another embodiment of the preset invention. In Fig. 7, the numeral 71 indicates band division means for dividing a voice signal containing noise for each frequency band as an example of frequency analysis means for frequency analyzing a signal, the numeral 72 does noise prediction means for inputting therein the output of the band division means 71 to predict a noise component, the numeral 73 does cancel means for eliminating the noise in a such a manner as described later, and the numeral 74 does band composition means for composing a voice as an example of signal composition means for composing a signal.

The above-described means will be explained in detail hereinafter. The band division means 71 is supplied with a voice containing noise/noise input, performs band division into m-channel frequency bands, and supplies them to the noise prediction means 72 and the cancel means 73. The noise prediction means 72 predicts noise component for each channel on the basis of the voice/noise input divided into m-channels, and supplies them to the cancel means 73. For example, the noise prediction is performed as described previously and as shown in Fig. 4. The cancel means 73 is supplied with a m-channel signal from the band division means 71 and the noise prediction means 72, cancels noise in a manner to subtract the noise for each channel in response to the cancel coefficient input, and supplies them to the band composition means 74. That is, the cancellation is performed by multiplying the predicted noise component by the cancel coefficient. Usually, the cancellation with the time axis as an example of cancel method is performed as

BNSDOCID: <EP_____0438174A2_I_>

described previously and as shown in Fig. 5. Also, as described previously and as shown in Fig. 6, the cancellation with the frequency as a basis is performed. The band composition means 74 composes the m-channel signal supplied from the cancel means 73 to obtain a voice output.

The operation of the signal processing device thus configured in the embodiments of the present invention will be explained hereinafter.

A voice containing noise/noise input is band divided into m-channel signals by the band division means 71, and a noise component thereof is predicted for each channel by the noise prediction means 72. From the signal band-divided into m-channels by the band division means 71, the noise component supplied for each channel from the noise prediction means 72 is eliminated. The noise elimination ratio at that time is properly set so as to improve articulation for each channel by the cancel coefficient input. For example, articulation is improved in such a manner that, where a voice signal exists, the cancel coefficient is made low even if a noise exists so as not to eliminate much the noise. Then, the noise-eliminated m-channel signal obtained by the cancel means 73 is composed by the band composition means 73 to obtain a voice output.

As described above, according to the present embodiment, the noise elimination ratio of the cancel means 73 can be properly set for each band by the cancel coefficient input, and the cancel coefficient is accurately selected according to a voice, thereby allowing an articulation noise-suppressed voice output to be obtained.

Fig. 8 is a block diagram of a signal processing device in an embodiment of another present invention.

In Fig. 8, the same numeral is assigned to the same means as that of the embodiment in Fig. 7. That is, the numeral 71 indicates band division means, the numeral 72 does noise prediction means, the numeral 73 does cancel means, and the numeral 74 does band composition means. Pitch frequency detection means 87 detects a pitch frequency of the voice of the inputted voice/noise input and supplies it to channel coefficient setting means 88. The pitch frequency of the voice is determined by various method as shown in Table 1 and expresses the presence/absence and characteristic of a voice.

Instead of the pitch frequency detection means 87, another means for detecting voice portion may also be provided. The cancel coefficient setting means 88 is configured in a manner to set a number m of cancel coefficients on the basis of the pitch frequency supplied from the pitch frequency detection means 87 and supply them to the cancel means 73.

The operation of the signal processing device thus configured in the embodiments of the present invention will be explained hereinafter.

A voice containing noise/noise input is band-divided into m-channel signals by the band division means 71, and a noise component thereof is predicted for each channel by the noise prediction means 72. From the signal band-divided into m-channels by the band division means 71, the noise component supplied for each channel from the noise

BNSDOCID: <EP 0438174A2_I_>

Table 1

Classification	Pitch extraction method	Features
I Waveform Processing	(1)Parallel Processing	Performs decision by majority of pitch periods extracted by 6-class simple waveform peak detector.
	(2)Data Reduction	Casts away data other than pitch pulse candidates from waveform data by various logic operations.
	(3)Zero Crossing Count	Aims at repeated pattern with respect to number of zero crosses of waveform.
II Correlation Processing	(4) Self-correlation	Self-correlation function of voice waveform and center clip allows spectrum to be smoothed and peak clip allows operation to be simplified.
	(5,a) Modified Correlation	Self-correlation function of residuals of LPC analysis, and LPF and polarization of residual signal allow operation to be simplified.
	(5,b) SIFT Algorithm	After down sampling voice waveform, performs LPC analysis and smoothes spectrum by inverse filter. Recovers time accuracy by interpolation of correlation function.
	(6)AMDF	Detects periodicity by amplitude mean difference function (AMDF). Also, allows residual signal to be extracted by AMDF.
III Spectrum Processing	(7)Cepstrum	Separates envelope and fine structure of spectrum by Fourier transformation of logarithm of power spectrum.
	(8)Period Histogram	Determines histogram of higher harmonic of basic frequency on spectrum, and decides pitch by common divisor of higher harmonics.

prediction means 72 is eliminated. The noise elimination ratio at that time is set for each channel by the cancel coefficient supplied from the cancel coefficient setting means 88. That is, when the predicted noise component represents a_i, signal containing noise b_i, and cancel coefficient $\underline{alpha}_i$, the output c_i of the cancel means 73 becomes (b_i - $\underline{alpha}_i$ x a_i). And the cancel coefficient thereof is determined on the basis of the information from the pitch frequency detection means 87. That is, the pitch frequency detection means 87 inputs therein a voice/noise input and detects the pitch frequency of the voice. The cancel coefficient setting means 88 sets cancel coefficients as shown in Fig. 9 . That is, Fig.9(a) shows cancel coefficients at each band, where the f_0-f_3 indicates the entire band of the voice/noise input. The f_0-f_3 is divided into m-channels to set the cancel coefficient. The f_1-f_2 indicates particularly a band containing voice obtained utilizing pitch frequency. Thus, in the voice band, the cancel coefficient is made low (close to zero) to eliminate noise as little as possible, thereby causing articulation to be improved. That is because human acoustic sense can hear a voice even though the voice has a little noise. In the non-voice bands f_0-f_1 and f_2-f_3, the cancel coefficient is made 1 to remove sufficiently noise. The cancel coefficient in Fig.9(b) is used when it is firmly found that no voice exists and only noise is considered to exist, and is made 1 to remove sufficiently noise.

BNSDOCID: <EP_____0438174A2_I_>

For example, where no vowel continues to exist from view of peak frequency, the signal cannot be determined to be voice signal, so that the signal is determined to be noise. It is preferable that the cancel coefficient in Fig.9 (a) and (b) can be properly changed over.

The present invention can apply not only to voice signal but also to other signals processing.

Also, the present invention, though implemented in software utilizing a computer, may also be implemented utilizing a dedicated hardware circuit.

As apparent by the above description, the signal processing device according to the present invention comprises noise prediction means for predicting a noise component, cancel means into which a noise-predicted output of the noise prediction means, a frequency analysis output of frequency analysis means and a cancel coefficient signal are inputted and which cancels the noise component considering the cancel ratio from the frequency analysis output, and signal composition means for composing the canceled output of the cancel means, so that, when the noise component is eliminated from a voice containing a noise, the degree of the elimination is properly controlled, thereby allowing the noise to be eliminated and articulation to be improved.

With reference to the drawings, an embodiment of another present invention will be explained.

Fig. 10 is a block diagram of a signal processing device in an embodiment of the present invention. The device is configured as shown in Fig. 10. That is, a noise prediction section 101 predicts a noise by a voice/noise input and by a control signal supplied by a voice detection section 103, and supplies a predicted noise to a cancel section 102. The cancel section 102 eliminates the noise from the voice/noise input in response to the predicted noise supplied from the noise prediction section 101 to obtain a voice output, and supplies the voice output to the voice detection section 103. The voice detection section 103 detects the presence/absence of actual voice by the voice output to obtain a voice-detected output, and supplies the voice-detected output as a control signal to the noise prediction section 101.

The operation in the above configuration will be explained. A voice overlapping with noise/noise input is supplied to the cancel section 102 where the noise is eliminated in response to the predicted noise supplied from the noise prediction section 101 to obtain a voice output. The voice/noise input from which the noise is eliminated by the cancel section 102 is supplied to the voice detection section 103 where the presence/absence of voice is detected to obtain a voice-detected output. Thus, the noise prediction section 101 operates such that the section uses as a control signal the voice-detected output indicating the presence/absence of a voice supplied from the voice detection section 103 to predict the noise of the voice/noise input signal, and supplies the voice-detected signal to the cancel section 102.

Thus, according to the signal processing device in the embodiment of the present invention, voice detection is performed by the signal in which a noise is previously eliminated from a voice/noise input, thereby allowing the presence/absence of a voice to be accurately detected regardless of noise. With such detection, noise prediction can also be performed accurately and the noise is eliminated effectively from the voice/noise input to obtain a clear voice output.

With reference to Fig. 11. an embodiment of another present invention will be explained hereinafter.

Fig. 11 is a block diagram of a signal processing device in an embodiment of the present invention. The device is configured as shown in Fig. 11. That is, a first cancel section 105 eliminates a noise predicted by a first noise prediction section 104 from a voice/noise input, and supplies the noise-eliminated signal to a voice detection section 106, a second noise prediction section 107 and a second cancel section 108. The voice detection section 106 detects the presence/absence of the signal supplied from the first cancel section 105 to obtain a voice-detected output, and supplies the voice-detected output as a control signal to the first noise prediction section 104 and the second noise prediction section 107. The second cancel section 108 eliminates the noise predicted by the second noise prediction section 107 from the signal supplied from the first cancel section 105 to obtain a voice output. The first noise prediction section 104 and the second noise prediction section 107 both use the control signal from the voice detection section 106 to predict the noise of the voice/noise input and to predict the noise of the signal supplied from the first cancel section 105, respectively. Then, the second noise prediction section 107 supplies the predicted-result to the second cancel section 108 which in turn makes the canceled-result a voice output.

The operation in the above configuration will be explained. A voice overlapping with noise/noise input is supplied to the first cancel section 105 where the noise is eliminated in response to a predicted noise supplied from the first noise prediction section 104. A first voice output from which the noise has been previously eliminated by the first cancel section 105 is supplied to the second cancel section 108 where the noise is further eliminated accurately in response to a second predicted noise supplied from the second noise prediction section 107 to obtain a voice output. Also, the first voice output from which the noise has been previously eliminated by the first cancel section 105 is supplied to the voice detection section 106 where the presence/absence is detected to obtain a voice-detected output (control signal). Now, the first

BNSDOCID: <EP_____0438174A2_I_>

noise prediction section 104 uses the control signal indicating the presence/absence of a voice supplied from the voice detection section 106 to predict the noise of the voice/noise input, and supplies a first noise-predicted signal to the first cancel section 105. Further, the second noise prediction section 107 operates such that the section 107 uses similarly the control signal indicating the presence/absence of a voice supplied from the voice detection section 106 to further predict accurately the noise from the first voice output signal from which the noise has been previously eliminated by the first cancel section 105, and supplies the second predicted noise to the second cancel section 108. Thus, according to the signal processing device in the second present invention, the presence/absence of a voice can be accurately detected regardless of noise, and the noise is further predicted accurately and eliminated from the first voice output from which the noise has been previously eliminated, thereby allowing a much lower level and rapidly fluctuated unsteady noise to be eliminated.

With reference to the drawing, an embodiment of another present invention will be explained hereinafter.

Fig. 12 is a block diagram of a signal processing device in an embodiment of another present invention.

The device is configured as shown in Fig. 12. That is, an FET processing section 121 transforms an input signal to a frequency-region signal, and supplies the transformed signal to a cepstrum peak detection section 122, a noise prediction section 125 and a cancel section 126. The cepstrum peak detection section 122 detects the cepstrum peak from the frequency-region signal obtained from the FET processing section 121, and supplies the detected cepstrum peak to a pitch frequency estimation section 123. The pitch frequency estimate section 123 estimates a pitch frequency from the cepstrum peak and supplies the pitch frequency to an window generation section 124 which in turn generates a window in response to the pitch frequency and supplies the window to a pitch frequency emphasis section 127. The noise prediction section 125 performs noise prediction for the signal supplied from the FET processing section 121 and supplies the noise-predicted signal to the cancel section 126 which in turn processes the signal supplied from the FET processing section 121 according to the predicted noise, and supplies the processed signal to the pitch frequency emphasis section 127. The pitch frequency emphasis section 127 performs pitch-frequency-emphasis-processing by the signals supplied from the window generation section 124 and the cancel section 126, and supplies the processed result to an IFFT section 128 which in turn transforms the signal to a time-region signal for output.

The operation in the above configuration will be explained. First, an input signal to the present device is transformed to a frequency-region signal by the FET processing section 121. The input signal transformed to frequency region is detected for the cepstrum peak thereof by the cepstrum peak detection section 122, and further determined for the pitch frequency thereof by the pitch frequency estimate section 123. Then, on the basis of the determined pitch frequency, the window generation section 124 generates a proper window to perform voice emphasis as the frequency-region data, and supplies the window to the pitch frequency emphasis section 127. The noise prediction section 125 performs noise prediction for the input signal transformed to frequency region, determines the noise component in the frequency region, and supplies the noise component to the cancel section 126. Then, the cancel section 126 eliminates accurately for each frequency component the noise component in the frequency region obtained by the noise prediction section 125 , from the input signal transformed to the frequency-region signal supplied from the FET processing section 121, and supplies the noise-eliminated signal to the pitch frequency emphasis section 127. Then, the pitch frequency emphasis section 127 controls the noise-eliminated frequency signal obtained from the cancel section 126 in response to the window to perform voice emphasis obtained from the window generation section 124, performs voice emphasis, and supplies the voice-emphasized signal to the IFFT processing section 128. Then, the IFFT processing section 128 transforms the signal from the pitch frequency emphasis section 127 to a time-region signal for output.

Thus, according to the signal processing device in the embodiment of the present invention, a noise is eliminated from the signal in which a voice overlaps the noise, and the pitch frequency emphasis section is provided to emphasize the voice component, thereby allowing a voice signal with an excellent articulation to be obtained.

Although it is apparently preferable that the window generated by the window generation section 124 in the above embodiment represents a voice harmonic wave structure, the window may be a comb filter and a low-pass filter. It is also apparent that the pitch frequency emphasis section 127 can be simply implemented in a multiplication circuit.

As apparent by the above embodiments, according to the present invention, a device which eliminates a noise by transforming a signal to frequency-region comprises pitch frequency prediction means for predicting a pitch frequency, window generation means for generating a window in response to the pitch frequency, noise prediction means, cancel means for eliminating the noise in response to the output of the

BNSDOCID: <EP_____0438174A2_I_>

noise prediction means, and pitch frequency emphasis means for emphasizing the pitch of the canceled-output of the cancel means using the window of the window generation means, whereby the noise can be eliminated from the signal in which a voice overlaps the noise and further the voice component be emphasized to obtain a voice signal with a high articulation.

It is further understood by those skilled in the art that the foregoing description is a preferred embodiment and that various changes and modificaitons may be made in the invention without departing from the spirit and scope thereof.

Claims

1. A signal processing device comprising:
frequency analysis means (3) for inputting therein a signal containing a noise to perform frequency analysis;
signal detection means (20) for detecting a signal portion from said frequency-analyzed signal ;
noise prediction means (6) for inputting therein said frequency-analyzed signal to predict the noise of the signal portion on the basis of a past noise information;
cancel means (7) for subtracting said predicted noise from said frequency-analyzed signal ; and
signal composition means (8) for performing signal composition on said noise-canceled signal.

2. A signal processing device in accordance with claim 1, wherein
the noise prediction means (6) predicts the noise of the signal portion using as a trigger the signal detected by said signal detection means (20).

3. A signal processing device in accordance with claim 1 or 2, wherein
said signal detection means (20) comprises :
cepstrum analysis means (4) for performing cepstrum analysis on the said frequency-analyzed signal; and
signal detecting means (5) for detecting the signal portion utilizing said analyzed cepstrum.

4. A signal processing device in accordance with claim 3, wherein
said signal detecting means (5)includes:
peak detection means (51) for the peak on said analyzed cepstrum; and
signal-noise detection means (52) for discriminating the signal on the basis of the detected peak information.

5. A signal processing device in accordance with claim 1, 2 or 3, wherein
said noise prediction means (6) accumulates the noise information utilizing the signal detected by said signal detection means (20)

6. A signal processing device comprising:
frequency analysis means (71) for analyzing the frequency of an inputted signal containing a noise;
noise prediction means (72) for inputting therein said frequency-analyzed output of the frequency analysis means (71) to predict the noise component thereof;
cancel means (73) for inputting therein said predicted noise of said noise prediction means (72), said frequency-analyzed output of said frequency analysis is means (71) and a cancel coefficient signal to cancel the noise component considering the cancel coefficient from said frequency-analyzed output; and
signal composition means (74) for composing the canceled-output of said cancel means (73).

7. A signal processing device comprising:
band division means (71) for dividing the frequency band of an inputted signal containing a noise;
noise prediction means (72) for inputting therein said band-divided output of said band division means (71) to predict the noise component;
cancel means (73) for inputting therein said predicted noise output of said noise prediction means (72), said band-divided output of said band division means (71) and a cancel coefficient signal to cancel said noise component considering the cancel coefficient from said band-divided output; and
band composition means (74) for composing said canceled-output of the cancel means.

BNSDOCID: <EP_____0438174A2_I_>

8. A signal processing device in accordance with claim 7, wherein
said device includes:
pitch frequency detection means (87) for detecting a pitch frequency of said input signal containing noise; and
cancel coefficient setting means (88) for setting said cancel coefficient signal in response to said pitch frequency output of said pitch frequency detection means (87)

9. A signal processing device comprising:
a noise prediction section (101) for outputting a predicted noise signal by a voice/noise signal input;
a cancel section (102) for inputting therein said voice/ noise signal input and said predicted noise signal to obtain a noise-eliminated voice output;
a voice detection section (103) for detecting the presence/absence of a voice signal from said noise-canceled voice signal to output a control signal corresponding to said presence/absence; and
said noise prediction section (101) performing noise prediction utilizing said control signal of said voice detection section.

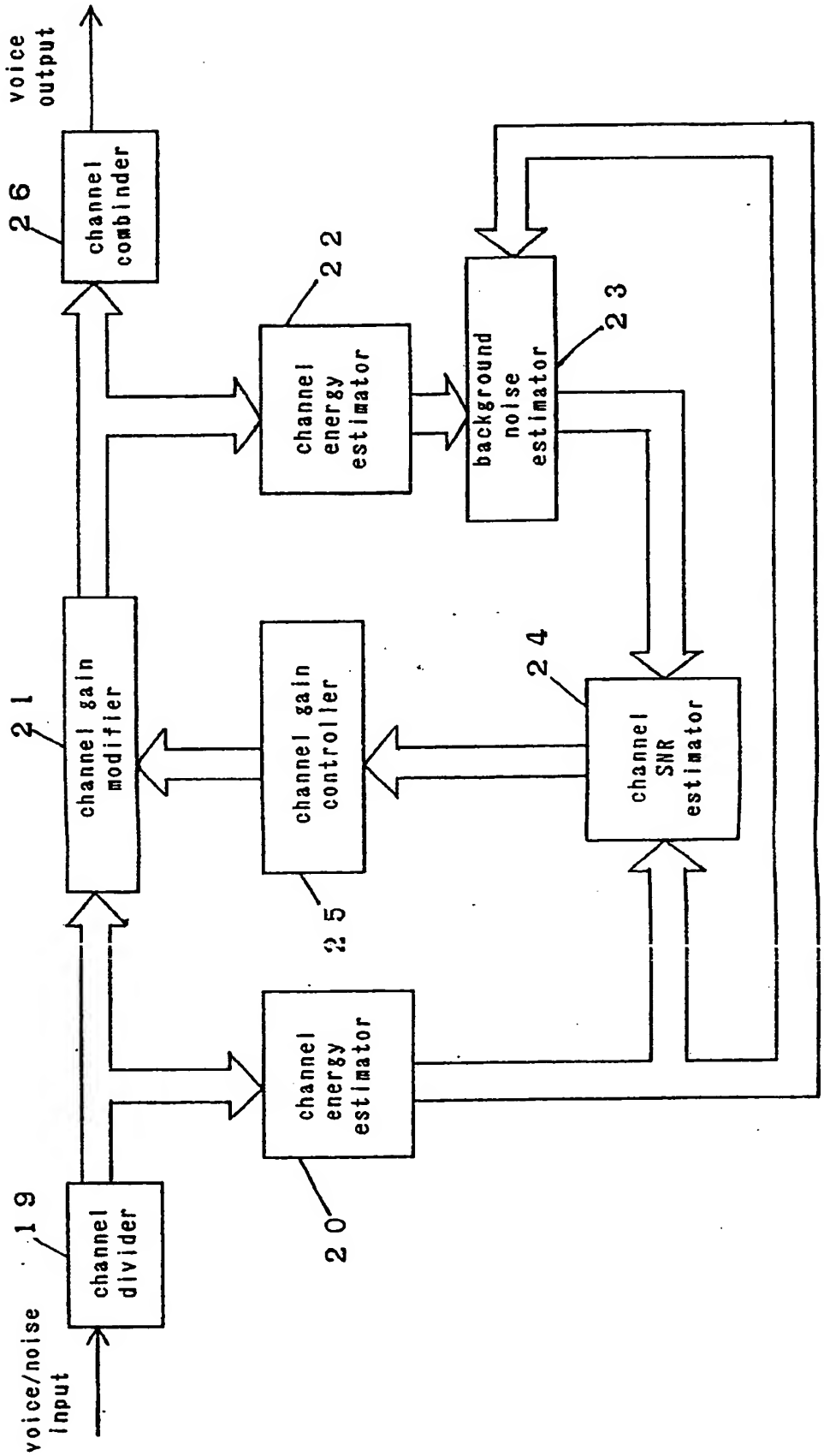
10. A signal processing device comprising:
a first noise prediction section (104) for outputting a first predicted noise signal by a voice/noise signal input;
a first cancel section (105) into which said voice/noise signal and said first predicted noise signal are inputted , to obtain a first noise-eliminated voice output;
a voice detection section (106) for detecting the presence/absence of a voice signal from said noise-canceled voice signal to output a control signal corresponding to the presence/absence;
said first noise prediction section (104) performing noise prediction utilizing said control signal of said voice detection section; and further
said device provided with ;
a second noise prediction section (107) for outputting a second predicted noise signal by said first voice output and by said control signal;
a second cancel section (108) into which said first noise-eliminated output and said second predicted noise signal are inputted to obtain a second noise-eliminated voice output.

11. A signal processing device comprising:
FFT processing means (121) for FFT processing a signal containing a noise;
cepstrum peak detection means (122) for performing cepstrum analysis of the FFT output of said FFT processing means (121) to detect the peak thereof;
pitch frequency estimate means (123) for estimating a pitch frequency in the peak-detected output of said cepstrum peak detection means (122);
window generation means (124) for generating a window output on the basis of the pitch frequency predicted output in said pitch frequency estimate means (123);
noise prediction means (125) for predicting a noise from the FFT output of said FFT processing means (121);
cancel means (126) for canceling the predicted noise output of said noise prediction means (125) from the FFT output of said FFT processing means (121);
pitch frequency emphasis means (127) for emphasizing the canceled output of said cancel means (126) by the window output of said window generation means (124); and
IFFT processing means (128) for IFFT processing the emphasized output of said pitch frequency emphasis means (127).

12. A signal processing method comprising the steps of:
analyzing the frequencies of an inputted signal containing noise;
detecting a signal portion from said frequency-analyzed signal;
predicting noise of the inputted frequency-analyzed signal portion on the basis of a past noise information;
subtracting said predicted noise from said frequency-analyzed signal; and
performing signal composition on said noise-cancelled signal.

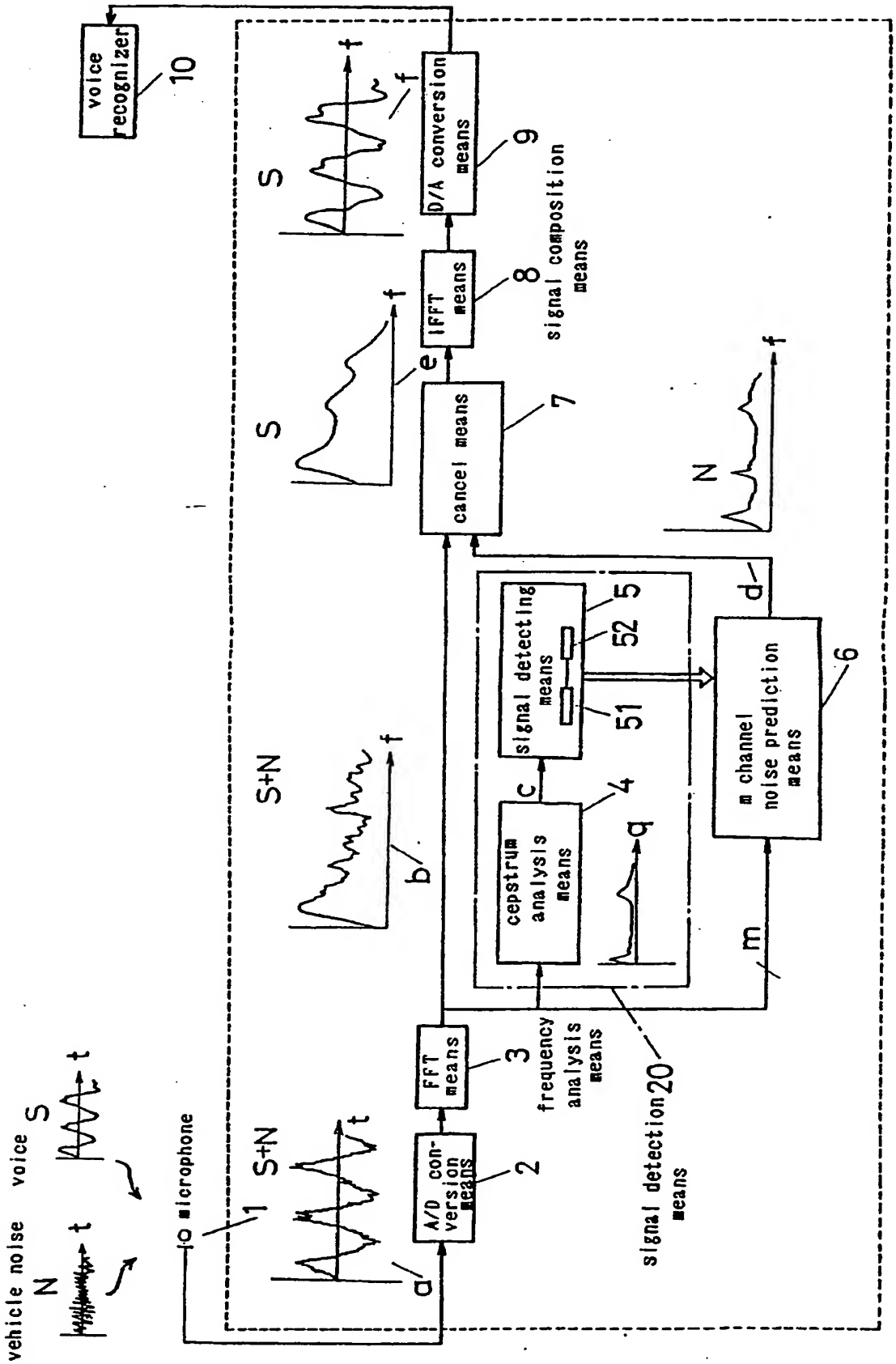
BNSDOCID: <EP_____0438174A2_I_>

FIG. 1 (PRIOR ART)



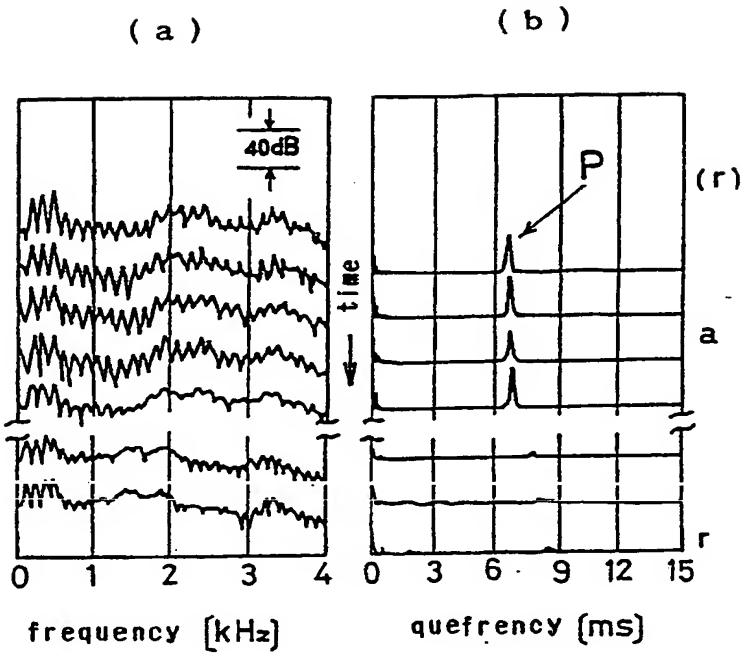
BNSDOCID: <EP_____0438174A2_I_>

FIG. 2



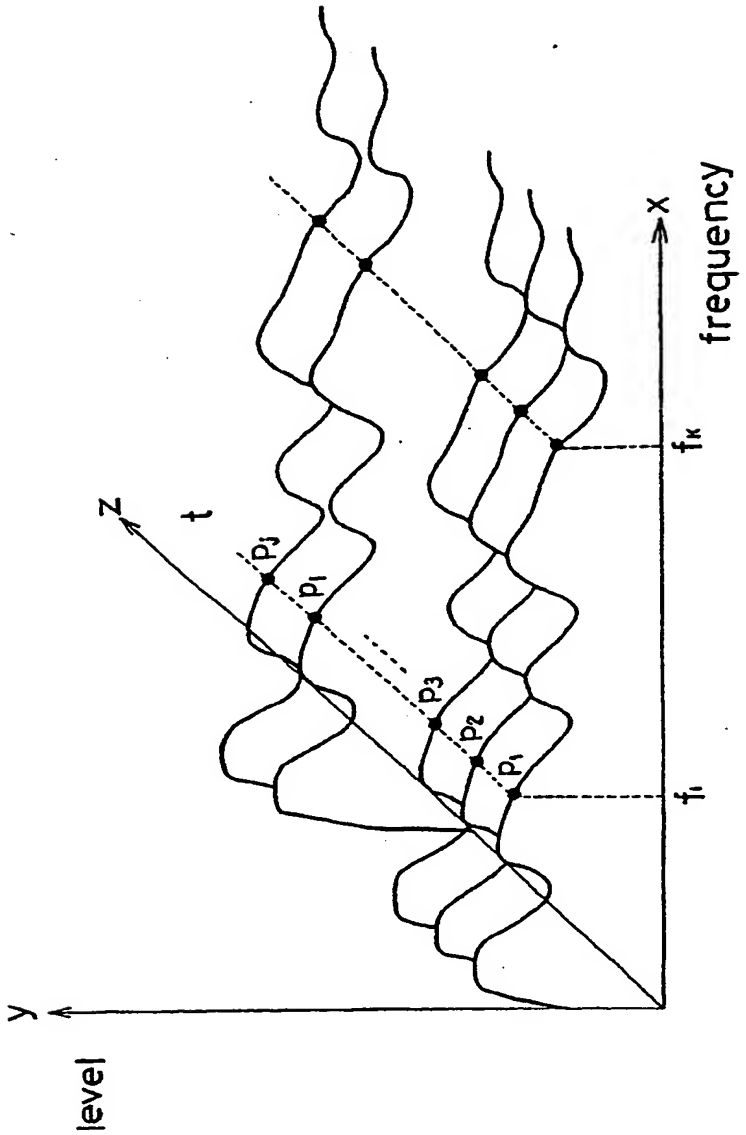
BNSDOCID: <EP_____0438174A2_I_>

FIG. 3



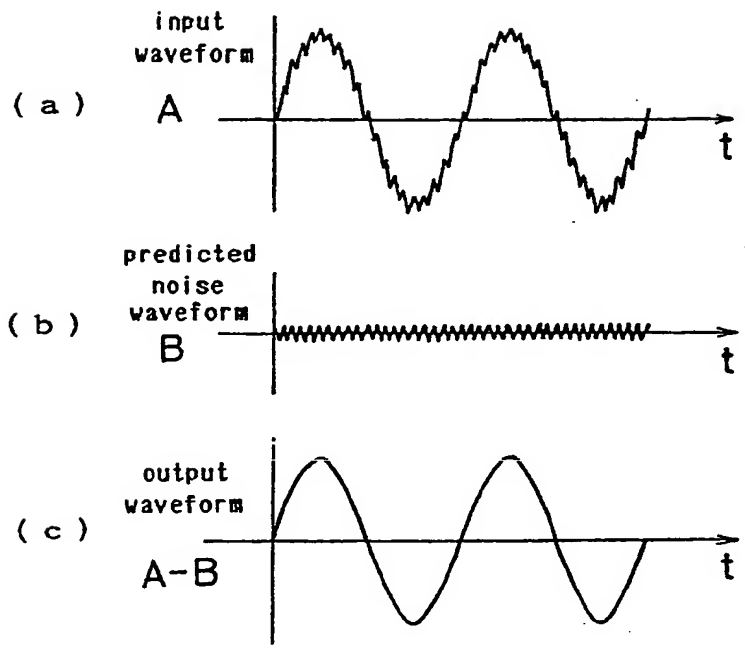
BNSDOCID: <EP_____0438174A2_I_>

FIG. 4



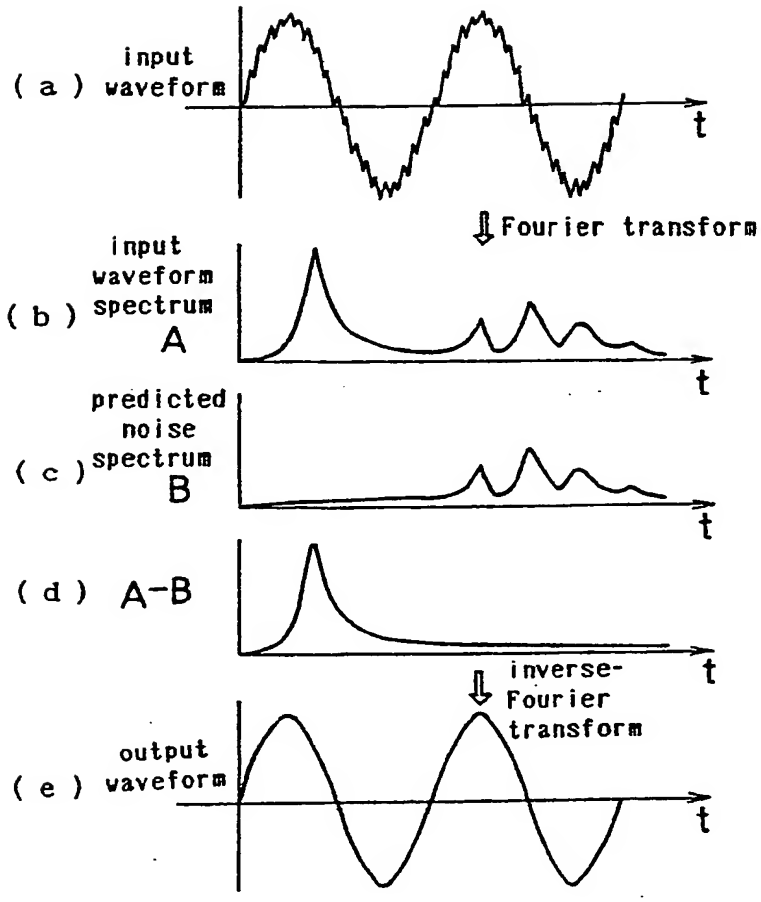
BNSDOCID: <EP_____0438174A2_I_>

FIG. 5



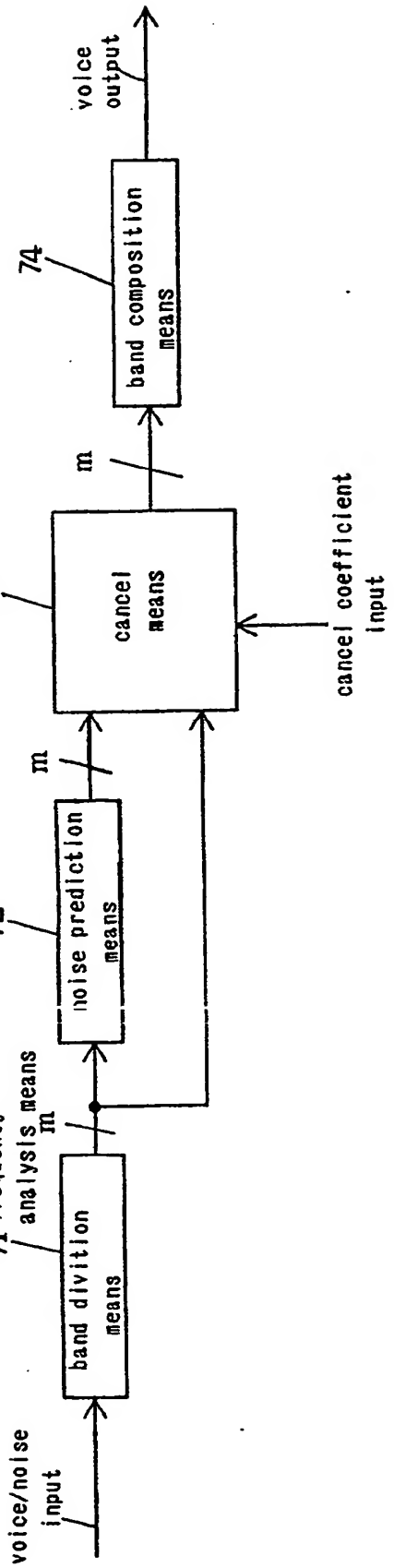
BNSDOCID: <EP_____0438174A2_I_>

FIG. 6



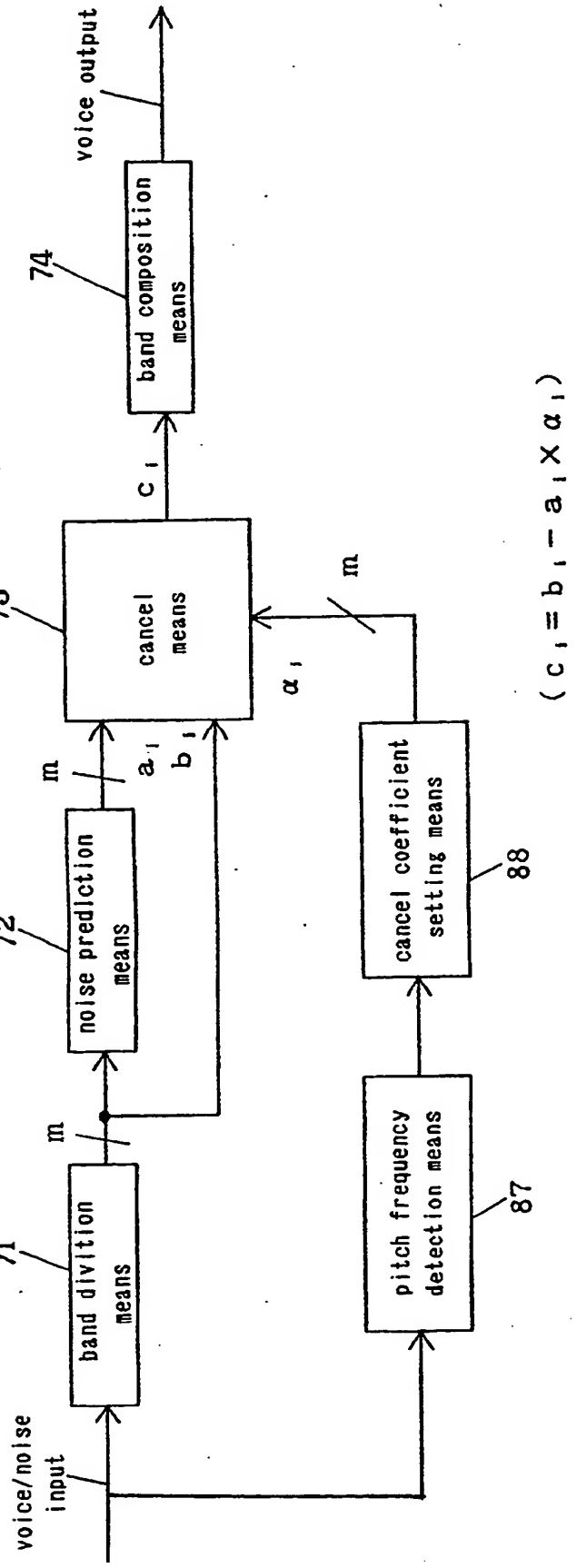
BNSDOCID: <EP_____0438174A2_I_>

FIG. 7



EPNSDOCID: <EP_____0438174A2_I_>

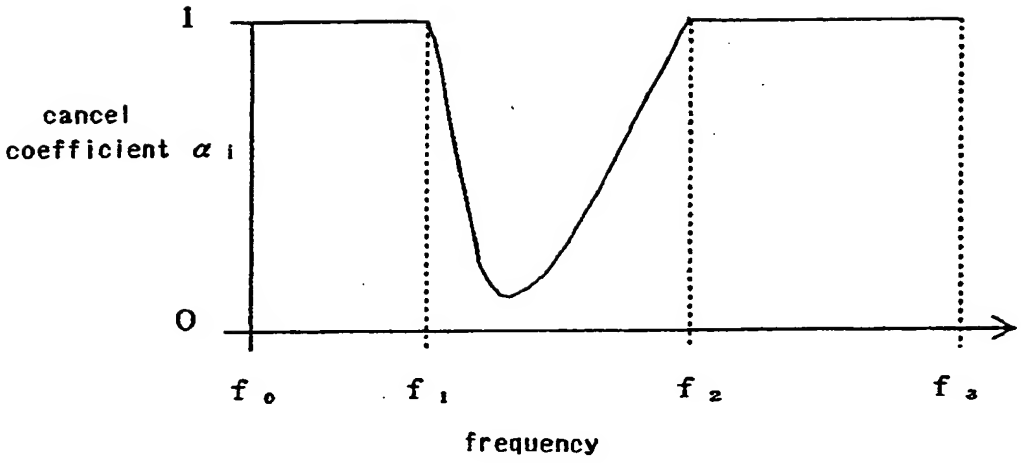
FIG. 8



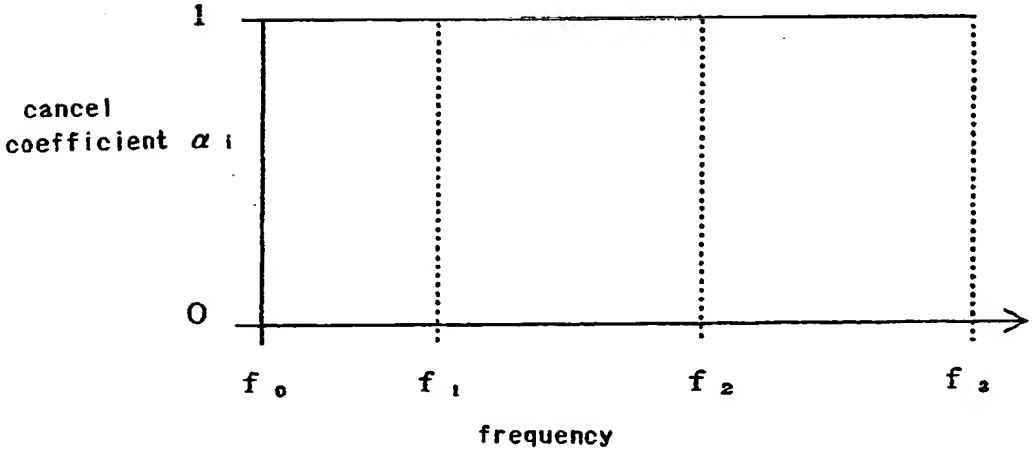
BNSDOCID: <EP_____0438174A2_I_>

FIG. 9

(a)

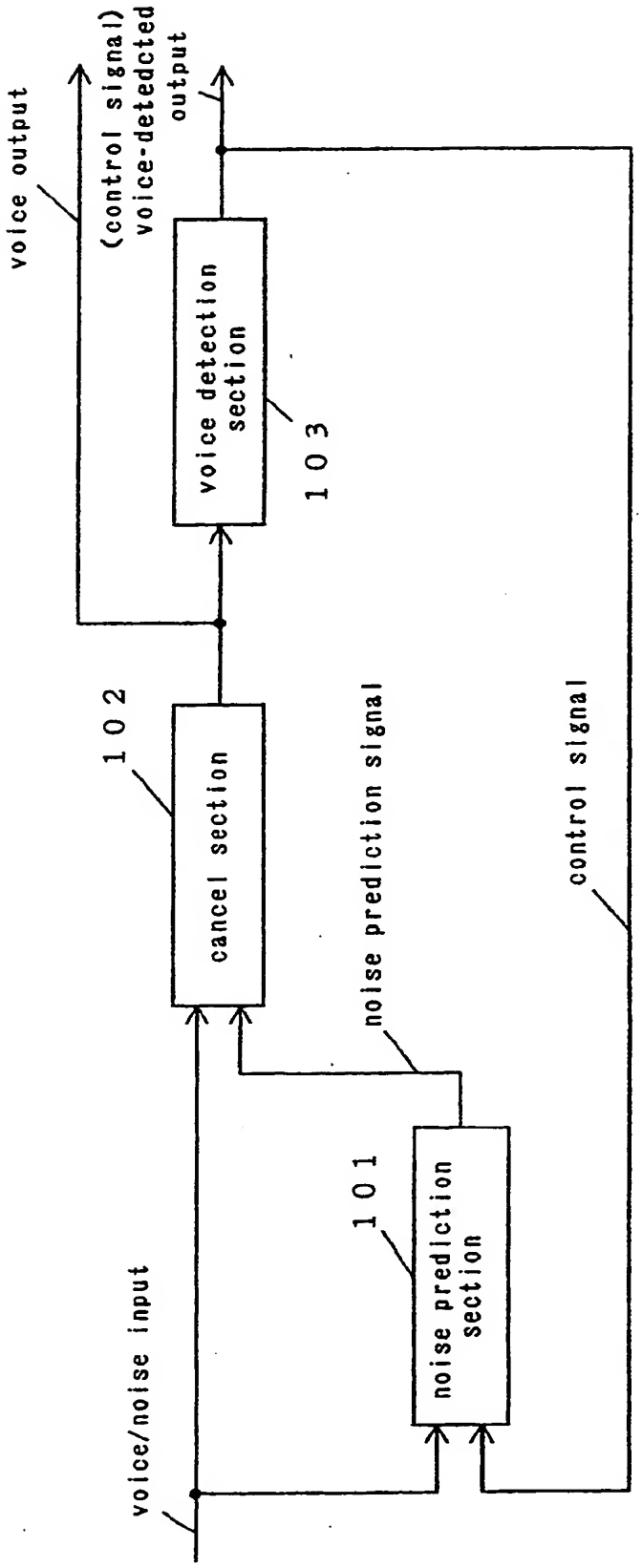


(b)



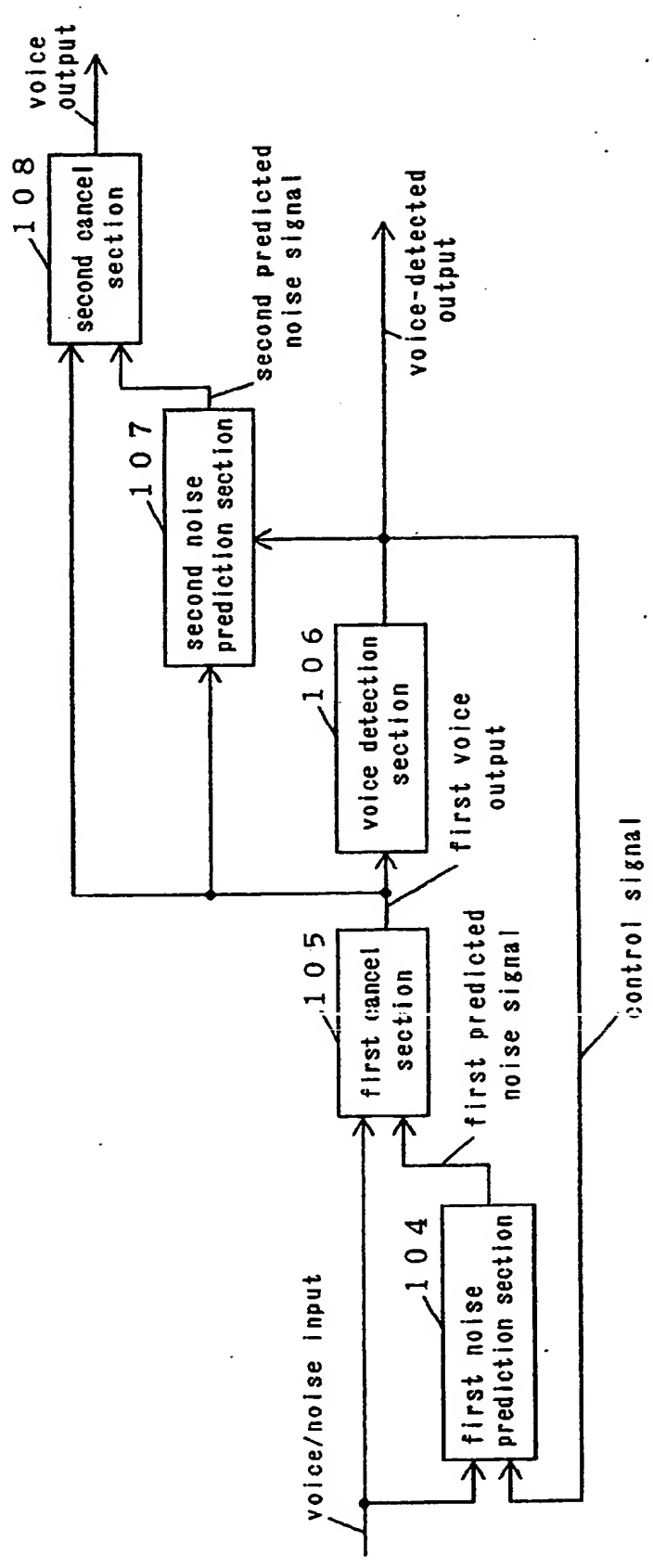
BNSDOCID: <EP_____0438174A2_I_>

FIG. 10



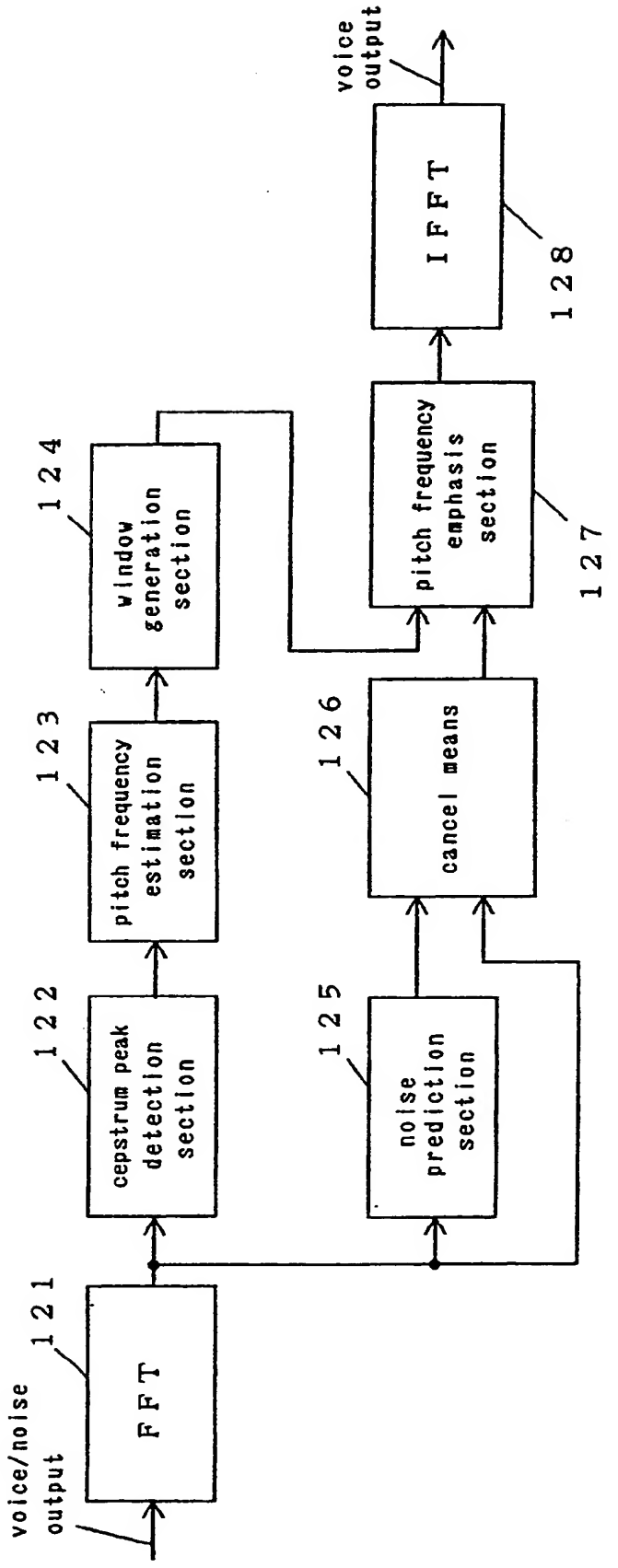
BNSDOCID: <EP_____0438174A2_I_>

FIG. 11



BNSDOCID: <EP_____0438174A2_I_>

FIG. 12



BNSDOCID: <EP_____0438174A2_I_>

(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: **0 438 174 A3**

(12) EUROPEAN PATENT APPLICATION

(21) Application number: 91100591.6

(51) Int. Cl.5: G10L 3/00

(22) Date of filing: 18.01.91

(30) Priority: 18.01.90 JP 8593/90
18.01.90 JP 8594/90
13.02.90 JP 33209/90
13.02.90 JP 33212/90

(43) Date of publication of application:
24.07.91 Bulletin 91/30

(84) Designated Contracting States:
CH DE FR GB LI NL SE

(88) Date of deferred publication of the search report:
11.09.91 Bulletin 91/37

(71) Applicant: MATSUSHITA ELECTRIC INDUSTRIAL CO., LTD.
1006, Oaza Kadoma
Kadoma-shi, Osaka-fu, 571(JP)

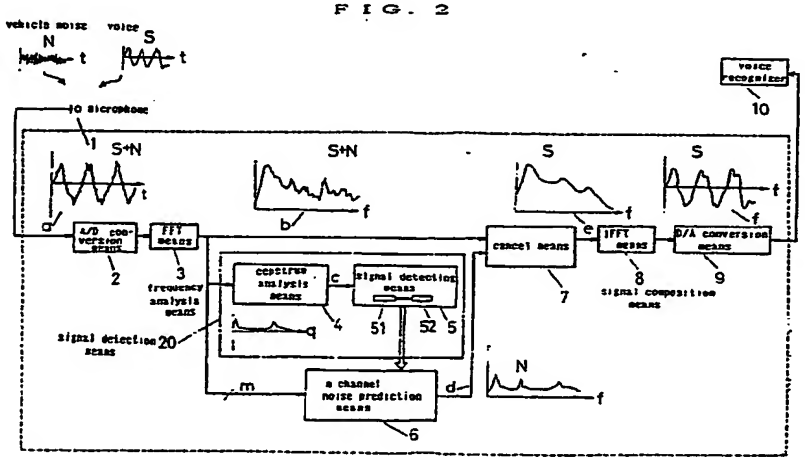
(72) Inventor: Kane, Joji
4-34, Jingu 1-chome
Nara-shi Nara 631(JP)
Inventor: Nohara, Akira
10-17, Matsuzonocho
Nishinomiya-shi Hyogo 662(JP)

(74) Representative: Patentanwälte Beetz sen. - Beetz jun. Timpe - Siegfried - Schmitt-Fumian- Mayr
Steinsdorfstrasse 10
W-8000 München 22(DE)

(54) Signal processing device.

(57) Frequency analysis means (3) inputs therein a signal containing a noise to perform frequency analysis; signal detection means (20) detect a signal portion from the frequency-analyzed signal ; noise prediction means (6) input therein the frequency-analyzed signal with a mingling of noise to predict a noise of the signal portion on the basis of the past information; cancel means (7) subtract the predicted noise from said frequency-analyzed signal with a mingling of noise; and band composition means (8) perform band composition on the canceled signal.

FIG. 2



EP 0 438 174 A3

BNSDOCID: <EP_____0438174A3_I_>

European Patent Office

EUROPEAN SEARCH REPORT

Application Number

EP 91 10 0591

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
X	SIGNAL PROCESSING, vol. 15, no. 1, July 1988, pages 43-56, Amsterdam, NL; N. DAL DEGAN et al.: "Acoustic noise analysis and speech enhancement techniques for mobile radio applications" * Pages 52-53, paragraph 4.4: "Spectral subtraction technique" *	1,2,5	G 10 L 3/00
Y	IDEM	3,4	
X	FREQUENZ, vol. 42, nos. 2-3, February-March 1988, pages 79-84, Berlin, DE; K. KROSCHEL: "Umgebungsgeräuschreduktion bei Sprachkommunikationssystemen" * Page 80, paragraph 2: "Geräuschunterdrückung durch spektrale Subtraktion"; Bild 1 *	1,2,5	
Y	INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, Dallas, Texas, 6th - 9th April 1987, pages 205-208; J.A. NAYLOR et al.: "Techniques for suppression of an interfering talker in co-channel speech" * Page 206: "Pitch estimation" *	3,4	TECHNICAL FIELDS SEARCHED (Int. Cl.5)
X	EP-A-0 076 687 (SIGNATRON) * Claims 1-12 *	6,7	G 10 L
A	ELECTRONICS & COMMUNICATIONS IN JAPAN, vol. 64, no. 9, September 1981, pages 21-30, Silver Spring, Maryland, US; H. NAGABUCHI et al.: "Quality improvement of synthesized speech in noisy speech analysis-synthesis processing" * Pages 23-26, paragraph 3: "Speech analysis synthesis considering additive noise"; figures 3,5 *	1,2,9	

The present search report has been drawn up for all claims

Place of search	Date of completion of search	Examiner
The Hague	18 June 91	FARASSOPOULOS A.

CATEGORY OF CITED DOCUMENTS

X: particularly relevant if taken alone
Y: particularly relevant if combined with another document of the same category
A: technological background
O: non-written disclosure
P: intermediate document
T: theory or principle underlying the invention

E: earlier patent document, but published on, or after the filing date
D: document cited in the application
L: document cited for other reasons

&: member of the same patent family, corresponding document

BNSDOCID: <EP 0438174A3 I >

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☑ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☑ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☑ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.